



Year: 2012

Validity and reproducibility of a physical activity questionnaire for older adults: questionnaire versus accelerometer for assessing physical activity in older adults

Siebeling, Lara ; Wiebers, Sarah ; Beem, Leo ; Puhan, Milo A ; Ter Riet, Gerben

Abstract: **BACKGROUND:** Physical activity (PA) is important in older adults for the maintenance of functional ability. Assessing PA may be difficult. Few PA questionnaires have been compared to activity monitors. We examined reproducibility and validity of the self-administered Longitudinal Ageing Study Amsterdam Physical Activity Questionnaire (LAPAQ) against a triaxial accelerometer (ACTR) (Sensewear®) Pro in older adults. **METHODS:** Participants wore the ACTR continuously for two weeks. After 2 (T [time] = 1) and 4 (T = 2) weeks, participants completed the LAPAQ. Since the LAPAQ asks about 2 weeks' worth of physical activity, the ACTR and LAPAQ coincided at T1. T2 was used to assess the reproducibility of the LAPAQ results only. We calculated Pearson's correlation coefficients (PCC) to examine reproducibility and validity. For visualization, we used scatterplots and Bland-Altman plots. With a receiver operating characteristics (ROC) curve we assessed how well the LAPAQ identifies older adults whose activity level is below official recommendations. **RESULTS:** A total of 89 persons were included. Of the participants, 48% were men; median age was 73, and median body mass index was 25. The 2-week mean total duration of activity was 2788 (ACTR, T = 1), 2439 (LAPAQ T = 1), and 1994 (LAPAQ T = 2) minutes. As a reference, 2 full weeks contained 20,160 minutes. Reproducibility of the LAPAQ was moderate (PCC 0.68, 95% CI 0.55-0.80). The median difference between LAPAQ at T = 1 and the ACTR (LAPAQ minus ACTR) was -510 minutes and the PCC was 0.25 (95% CI 0.07-0.44). The area under the ROC curve was 0.73 (95% CI 0.59-0.86). **CONCLUSION:** LAPAQ underestimates PA and seems unsuitable for exact measurement in older adults. However, it may be used to determine if a person's PA level is below the recommended level.

DOI: <https://doi.org/10.2147/CLEP.S30848>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-74566>

Journal Article

Published Version

Originally published at:

Siebeling, Lara; Wiebers, Sarah; Beem, Leo; Puhan, Milo A; Ter Riet, Gerben (2012). Validity and reproducibility of a physical activity questionnaire for older adults: questionnaire versus accelerometer for assessing physical activity in older adults. *Clinical Epidemiology*, 4:171-80.

DOI: <https://doi.org/10.2147/CLEP.S30848>

Validity and reproducibility of a physical activity questionnaire for older adults: questionnaire versus accelerometer for assessing physical activity in older adults

Lara Siebeling¹
Sarah Wiebers²
Leo Beem¹
Milo A Puhan³
Gerben ter Riet¹

¹Department of General Practice,

²Faculty of Medicine, Academic Medical Centre, University of Amsterdam, Amsterdam, the Netherlands; ³Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Background: Physical activity (PA) is important in older adults for the maintenance of functional ability. Assessing PA may be difficult. Few PA questionnaires have been compared to activity monitors. We examined reproducibility and validity of the self-administered Longitudinal Ageing Study Amsterdam Physical Activity Questionnaire (LAPAQ) against a triaxial accelerometer (ACTR) (Sensewear® Pro) in older adults.

Methods: Participants wore the ACTR continuously for two weeks. After 2 (T [time] = 1) and 4 (T = 2) weeks, participants completed the LAPAQ. Since the LAPAQ asks about 2 weeks' worth of physical activity, the ACTR and LAPAQ coincided at T1. T2 was used to assess the reproducibility of the LAPAQ results only. We calculated Pearson's correlation coefficients (PCC) to examine reproducibility and validity. For visualization, we used scatterplots and Bland–Altman plots. With a receiver operating characteristics (ROC) curve we assessed how well the LAPAQ identifies older adults whose activity level is below official recommendations.

Results: A total of 89 persons were included. Of the participants, 48% were men; median age was 73, and median body mass index was 25. The 2-week mean total duration of activity was 2788 (ACTR, T = 1), 2439 (LAPAQ T = 1), and 1994 (LAPAQ T = 2) minutes. As a reference, 2 full weeks contained 20,160 minutes. Reproducibility of the LAPAQ was moderate (PCC 0.68, 95% CI 0.55–0.80). The median difference between LAPAQ at T = 1 and the ACTR (LAPAQ minus ACTR) was –510 minutes and the PCC was 0.25 (95% CI 0.07–0.44). The area under the ROC curve was 0.73 (95% CI 0.59–0.86).

Conclusion: LAPAQ underestimates PA and seems unsuitable for exact measurement in older adults. However, it may be used to determine if a person's PA level is below the recommended level.

Keywords: physical activity, elderly, validation, questionnaire, accelerometer

Background

Physical activity (PA) is important in maintaining health and functional ability, especially in older adults. It may also be an important predictor for the course of chronic diseases such as chronic obstructive pulmonary disease.^{1–3} Several studies showed that lack of PA is a risk factor for the development of many chronic diseases.^{4–6} Regular PA can play an important role in the prevention and management of cardiovascular disease, hypertension, diabetes, and other chronic diseases.⁴ Warburton et al found that there is strong evidence that regular PA is effective in the primary and secondary prevention of several chronic diseases and premature death.⁷ The American College

Correspondence: L Siebeling
Academic Medical Centre, University of Amsterdam, Department of Clinical Methods and Public Health, Room J2-114, Meibergdreef 9, 1105 AZ Nederland
Tel +31 205 668 382
Email l.siebeling@amc.uva.nl

of Sports Medicine (ACSM) and The American Heart Association (AHA) have set PA guidelines. The basic ACSM and AHA recommendations for people over 65 years of age include moderate exercise for at least 30 minutes a day, five days a week, or vigorous exercise for at least 20 minutes a day, three days a week in combination with strength training exercises.⁸

However, the assessment of PA remains difficult, especially in older adults.^{9–11} A number of questionnaires have been validated to assess PA in older adults, but all have several limitations. For example, the Zutphen Physical Activity Questionnaire does not include household activities, one of the main activity types performed by older adults.¹² The Modified Baecke Questionnaire for Older Adults does not include walking and bicycling, which are common daily activities.¹³ The International Physical Activity Questionnaire includes walking and bicycling for transportation purposes, and also includes household activities, but these items have only been validated among 18–65-year-old adults.¹⁴ The Longitudinal Ageing Study Amsterdam Physical Activity Questionnaire (LAPAQ) was developed and validated by keeping these limitations in mind, but it nevertheless has some limitations itself.¹⁵ First, the questionnaire is interviewer-administered, which requires training and substantial resources for its application in practice and studies. A self-administered LAPAQ was recently developed, but has not undergone a validation process yet. Second, the LAPAQ was compared with a diary and pedometer, which were used as validation instruments. Neither the diary nor the pedometer was able to accurately measure PA or validate the findings of the LAPAQ. Third, one study investigated the extent to which the findings of the LAPAQ were reproducible.¹⁵ The questionnaire was administered twice, 1 year apart. One year is probably too long to assess test–retest reliability because a person's PA pattern can change substantially in 1 year. Nevertheless, if the LAPAQ turns out to have good measurement properties, it would facilitate practical PA assessment in older adults. At face value, the LAPAQ appears to be a promising tool for measuring PA; however, uncertainty about its measurement properties remains. Therefore, we examined the reproducibility and validity of the self-administered LAPAQ using a modern triaxial accelerometer as a validation instrument.^{16–19}

Methods

Setting, participants, and design

The target participants were all persons aged ≥ 65 years from one primary health care center registered in the research network of the Department of General Practice of the University

of Amsterdam in the Netherlands. These individuals were identified through electronic patient charts by general practitioners (GPs). Around 850 persons aged 65 years or older received study information and a written invitation letter from their GP to participate in the study. They were invited to return a reply card only if they were interested in participating in the study. Approximately 150 individuals indicated their interest in participating in the study, and we contacted them by phone, inviting them for a first visit. For organizational reasons, we could only include a maximum of 100 participants. Inclusion criteria were Dutch language as native language and the ability to walk independently, with or without assistive devices. Exclusion criteria were dementia, psychosis, or other psychiatric comorbidities that may invalidate the assessment of self-reported parameters such as the LAPAQ. These criteria were evaluated by GPs before invitations were sent and checked by study personnel at the first visit. During the first visit (T_0) all potential participants received study information and written informed consent was obtained. Patient characteristics such as age, sex, body mass index (BMI, weight [kg] /height [m²]), and information about comorbidities were collected. Elaborate instructions for wearing the accelerometer (ACTR) were given.

Participants wore the ACTR continuously for 2 weeks. After 2 ($T = 1$) and 4 ($T = 2$) weeks, participants completed the LAPAQ. Since the LAPAQ asks about participants' PA over the course of 2 weeks, the ACTR and the LAPAQ coincided at T_1 , while T_2 was used to determine the reproducibility of the LAPAQ only. Appendix 1 contains a flow chart of the study. The study was approved by the Medical Ethics Committee of the Academic Medical Center and was funded by the Dutch Asthma Foundation.

LAPAQ

The LAPAQ contains 18 questions, which cover the frequency and duration of six activities during the previous 2 weeks: walking outside, bicycling, gardening, light and heavy household activities, and sports activities. The participants were also asked if their physical activity pattern in the previous 2 weeks had been representative of the rest of the year. Stel and colleagues validated the interviewer-administered LAPAQ and adjusted this original version to obtain a less time-consuming and more practical self-administered version.¹⁵ First, some sentences were reformulated to improve participant comprehension. Second, only the nine most common sports activities were mentioned in the new version. Third, questions about moving around in a wheelchair were omitted. Finally, the last question ("were the previous two weeks representative for the rest

of the year” with answer possibilities “Yes” and “No, I did less because...” received an additional answer possibility (“No, I did more because...”). No changes with respect to the content of the questions were made and it is unlikely that the activity measurements are affected by the adjustments. Appendix 2 shows the self-administered LAPAQ that was used in this study, which takes 5–10 minutes to complete. The original LAPAQ is available in Dutch, German, and English. The total scores can be measured for each activity and summing up the scores across all activities provides a total PA duration score (in minutes/2 weeks). The intensity for each activity can be expressed as metabolic equivalent tasks (METs). One MET is defined as the ratio of work metabolic rate to a standard resting metabolic rate of 1.0 kilocalories per kilogram of body weight per hour.^{20,21} To compare the LAPAQ with the accelerometer, PA was classified into three MET-categories: 2–2.99 METs (mild PA), 3–5.99 METs (moderate PA), and ≥ 6 METs (vigorous PA). We translated the LAPAQ-reported PA into MET-values using a compendium^{20,21} (see Table 1).

According to the ACSM guidelines, cut-off scores for the different PA intensity levels are: 0–3 METs, mild PA; 3–6 METs, moderate PA; 6–9 METs, vigorous PA; and 9–12 METs, very vigorous PA. We used slightly different cut-off points because the PA intensity of the LAPAQ activities ranged from 2.5 to 8 METs.

Validation instrument

We used the Sensewear[®] Pro Armband Accelerometer (Body-Media, Inc, Pittsburgh, PA) as the reference point; this is a device that was validated with doubly labeled water and

indirect calorimetry.^{16–19} It is an ambulant body monitor system, which records metabolic and physical activity continuously. It gives an exact overview of energy expenditure, and the duration (minutes) and intensity of daily activities (MET scores). The device is attached to the right upper arm with a band, is $88 \times 56 \times 21$ mm, and weighs 82 grams. Participants wore the device 24 hours a day (except when showering or swimming) for 2 weeks. The mean time that the device was actually worn was calculated. The software provided information about the time spent in the above mentioned MET categories.

Statistical analysis

Many studies use correlation coefficients (CC) for the validation of questionnaires, but limitations of correlation coefficients are well documented in the literature.²² For example, since CCs are dependent on the true between-subject variation in the given study population, extrapolating results to other populations can be misleading. Furthermore, since CCs are measures of association but not of agreement between a questionnaire and its reference criterion, CCs cannot detect systematic errors. We calculated correlation coefficients, but our primary interest was a comparison by graphically visualizing aspects of agreement in scatter plots and Bland–Altman plots.²³ Differences in scores and Pearson CCs (PCCs) between LAPAQ scores and accelerometer values (validity), as well as differences in scores between the two LAPAQ measurements (reproducibility) were calculated for the three PA intensity categories (2–2.99 METs, 3–5.99 METs, and ≥ 6 METs). Usually, intraclass correlation coefficients (ICC) are used for measuring reproducibility; however, we used the PCC, which is related. One major difference between these two measurements is that the ICC centers around and is scaled by a pooled mean and standard deviation (SD), while the PCC centers around and is scaled by the mean and SD of each variable. Since the activity pattern may have been different in the second time period, we preferred to use the PCC. We used logistic regression and the area under the receiver operating characteristics (ROC) curve to assess how well the LAPAQ score discriminates between individuals whose activity level is or is not in accordance with the ACSM and AHA recommendations. All statistical analyses were performed with Stata/SE 10.1 (StataCorp LP, College Station, TX) and SAS 9.2 (SAS Institute Inc, Cary, NC).

Results

Subject characteristics

During September and October 2010, a total of 92 subjects were recruited. Of these participants, three were excluded

Table 1 Intensity levels of LAPAQ activities

Activity	Intensity (METs)	MET category
Walking	2.5	2–2.99
General bicycling	4.0	3–5.99
Gardening	4.0	3–5.99
Sports		
Gymnastics	4.0	3–5.99
Hometrainer cycling	7.0	≥ 6
Cycling tour	8.0	≥ 6
Walking tour	6.0	≥ 6
Swimming	6.0	≥ 6
Badminton/tennis	7.0	≥ 6
Winter sports	7.0	≥ 6
Cardio-fitness	5.5	3–5.99
Light household activities	2.5	2–2.99
Heavy household activities	4.0	3–5.99

Abbreviations: LAPAQ, Longitudinal Ageing Study Amsterdam Physical Activity Questionnaire; MET, metabolic exercise equivalent.

because they provided extremely high PA scores on the LAPAQ of 9170 minutes and 14,280 minutes (highly improbable given that there are 20,160 minutes in two weeks); one participant even provided a score of 22,680 minutes. In addition, one LAPAQ at T = 1 was missing, and two LAPAQs at T = 2 were missing. Thus, we had 86 complete records. Of the 86 subjects included in the study, 48% were male; median age was 72 years, and median BMI was 25. Participants actually wore the accelerometer on their body for 98.7 percent of the time (interquartile range from 97.8% to 99.2%). This percentage was measured as the actual on-body time divided by the theoretically maximal on-body time, which is 14 days. Table 2 shows subjects' characteristics and includes their chronic disease profile. A total of 36% participants did not report having any diseases; 43% reported having one disease; and 21% reported living with two or more diseases. Cardiovascular disease was reported most often (35%).

Table 3 shows mean, median, and interquartile range of PA scores in each MET category for the accelerometer, LAPAQ at T = 1 and LAPAQ at T = 2. As a reference, two full weeks contain 20,160 minutes (336 hours). According to the accelerometer, our population spent around 14% of their time (2748/20,160 minutes) on PA and 10% (2058/20,160 minutes) of their time was spent engaging in mild PA (walking and light household activities).

Reproducibility

For reproducibility analyses, LAPAQ scores at T = 1 were compared with LAPAQ scores at T = 2 (n = 86). PCCs, mean,

Table 2 Subject characteristics (n = 89)

Sex	
Male	43 (48.3)
Age	
Median, range	72.4, 65.4–87.6
BMI	
Median, range	25.0, 17.0–35.7
Number of diseases	
0	32 (35.9)
1	38 (42.7)
≥2	19 (21.3)
Cardiovascular disease	31 (34.8)
Musculoskeletal disease	17 (19.1)
Diabetes mellitus	6 (6.7)
Neurologic; CVA/TIA	4 (4.5)
Respiratory disease	4 (4.5)
Malignancy	1 (1.1)
Other disease	14 (15.7)

Note: Unless otherwise specified, numbers are absolute numbers and percentages, n (%).

Abbreviations: n, number; BMI, body mass index (weight [kg]/height[m]²); CVA, cerebrovascular accident; TIA, transient ischemic attack.

Table 3 Descriptive statistics for duration of physical activity in minutes in 2 weeks

	Accelerometer	LAPAQ at T = 1	LAPAQ at T = 2
Number (n)	89	88	87
2–2.99 METs			
Mean (SD)	1892 (827)	1630 (1194)	1299 (1070)
Median	2058	1275	1050
p25/p75	1325/2433	720/2520	420/2000
3–5.99 METs			
Mean (SD)	865 (665)	630 (740)	539 (563)
Median	663	440	440
p25/p75	343/1178	105/840	60/850
≥6 METs			
Mean (SD)	32 (119)	180 (386)	160 (298)
Median	3	0	0
p25/p75	0/15	0/170	0/210
All METs (≥2)			
Mean (SD)	2788 (1265)	2439 (1678)	1994 (1367)
Median	2748	1945	1760
p25/p75	1902/3713	1230/3540	890/2820

Notes: On average, scores were higher for accelerometer than for LAPAQ at T = 1 and for LAPAQ at T = 1 than for LAPAQ at T = 2. As a reference, 2 full weeks contain 20,160 minutes.

Abbreviations: n, number; LAPAQ, Longitudinal Ageing Study Amsterdam Physical Activity Questionnaire; T = 1, 2 weeks; T = 2, 4 weeks; MET, metabolic exercise equivalent; SD, standard deviation; p25, 25th percentile; p75, 75th percentile.

and median difference scores for all categories are shown in Table 4. The PCC for total PA (≥2 METs) was 0.68 (95% CI 0.55–0.80) and the mean and median difference (LAPAQ T = 1 minus LAPAQ T = 2) was 436 and 248 minutes, respectively.

PCCs were also calculated for the “representative” group (n = 50); these were the persons who claimed that their PA patterns were stable (2 × “yes” on question 18 [“were the previous two weeks representative for the rest of the year?”]). For the total group (≥2 METs), the PCC was 0.73 (95% CI 0.59–0.88) and for the 2–2.99 METs, 3–5.99 METs, and ≥6 METs, the PCCs were 0.69 (95% CI 0.54–0.84), 0.81 (95% CI 0.69–0.93) and 0.81 (0.49–0.93), respectively. In the scatter plot in Figure 1, the regression line is less steep than the line of equality. In general, average LAPAQ scores at T = 1 are higher than at T = 2. The Bland-Altman plot shows again that the higher the PA score, the larger the difference in scores between LAPAQ at T = 1 and T = 2 (see Figure 1).

Validity

For validity analyses, accelerometer scores were compared with scores on LAPAQ at T = 1 (n = 88). PCCs, mean, and median difference scores for all categories are shown in Table 4. The PCC was 0.25 (95% CI 0.07–0.44) for total PA (≥2 METs), and the mean and median difference in

Table 4 PCC and difference in scores for all MET categories

	2–2.99 METs	3–5.99 METs	≥6 METs	All METs (≥2)
PCC LAPAQ T = 1/T = 2 (95% CI)	0.58 (0.42–0.72)	0.79 (0.69–0.88)	0.75 (0.47–0.87)	0.68 (0.55–0.80)
Mean difference score (SD) (LAPAQ T = 1 minus LAPAQ T = 2)	309 (1004)	102 (436)	23 (258)	436 (1260)
Median difference score (LAPAQ T = 1 minus LAPAQ T = 2)	70	0	0	248
PCC LAPAQ T = 1/accelerometer (95% CI)	0.05 (–0.16–0.24)	0.27 (0.07–0.48)	0.01 (–0.07–0.25)	0.25 (0.07–0.44)
Mean difference score (SD) (LAPAQ T = 1 minus accelerometer)	–267 (1423)	–234 (852)	148 (403)	–354 (1830)
Median difference score (LAPAQ T = 1 minus accelerometer)	–363	–277	0	–510

Abbreviations: PCC, Pearson correlation coefficient; MET, metabolic exercise equivalent; LAPAQ, Longitudinal Ageing Study Amsterdam Physical Activity Questionnaire; SD, standard deviation; T = 1, 2 weeks; T = 2, 4 weeks; CI, confidence interval.

scores (LAPAQ T = 1 minus accelerometer) was –354 and –510 minutes, respectively. In the scatter plot, the regression line is less steep than the line of equality. Below 2000 minutes of PA, the average LAPAQ scores are higher than average accelerometer scores; when the scores are greater than approximately 2000 minutes of PA, the LAPAQ scores were lower than the average accelerometer scores. The Bland–Altman plot shows much spread around the regression line and the range between the limits of agreement is wide, indicating measurement error in the LAPAQ (see Figure 2).

The positive predictive value of the LAPAQ (the ability for this questionnaire to correctly predict whether PA levels are above the recommended level) is 0.88 (46/52). The negative predictive value of the LAPAQ (the ability of this questionnaire to correctly predict whether PA levels are below the recommended level) is 0.33 (12/36). The area under the ROC curve was 0.73 (95% CI 0.59–0.86).

Discussion

Due to moderate reproducibility and low validity, the LAPAQ seems unsuitable in providing an exact estimate of physical activity levels in older adults. Overall, the LAPAQ underestimates physical activity by 510 minutes (8.5 hours in two weeks, 36 minutes per day). However, for the highest MET-category (sports activities) the LAPAQ scores were higher than the scores provided by the accelerometer. This may be explained by a simple example: when a person is playing tennis for 1 hour, (s)he might fill in “1 hour” on the LAPAQ questionnaire, but the actual time engaging in PA above 6 METs is obviously less than 1 hour. When we compared both LAPAQ measurements, scores at T = 1 were higher than at T = 2, with a median difference of 248 minutes

(4 hours in 2 weeks, 18 minutes per day). This may be a systematic effect caused by participants being more active in the first 2 weeks because they knew that their activity was being measured by the accelerometer (otherwise known as the Hawthorne effect).²⁴ To avoid this, participants could have worn the accelerometer for 4 weeks instead of 2 weeks, but this might have led to a high drop-out rate since most participants were happy to take off the ACTR after two weeks. If this explanation is true, the reproducibility of the LAPAQ might be better than we measured in the current study.

Although the LAPAQ seems unsuitable for providing exact measurements of physical activity levels in older adults, a more modest aim of this questionnaire might be to use it to determine if a person's activity level is above the current ACSM and AHA recommendations that individuals engage in moderate exercise for at least 30 minutes a day, 5 days a week (300 minutes ≥3 METs in two weeks), or that individuals engage in vigorous exercise for at least 20 minutes a day, for 3 days a week (120 minutes ≥6 METs in two weeks). For this determination, the positive predictive value of LAPAQ is 88% and the negative predictive value is 33%. This means that LAPAQ incorrectly predicts PA levels above the recommended guidelines in only 12% of respondents.

PA is important for older adults to maintain their health and functional ability. A number of questionnaires have been validated to assess physical activity in older adults, but all of them have several limitations and most have not been compared to activity monitors. We compared the LAPAQ with Sensewear, an ACTR that was validated with doubly labeled water and indirect calorimetry.^{16,18} A second strength of our study is that the ACTR and LAPAQ measurements coincided at T = 1, implying that the LAPAQ covered the same time interval as the ACTR. With regards to reproducibility,

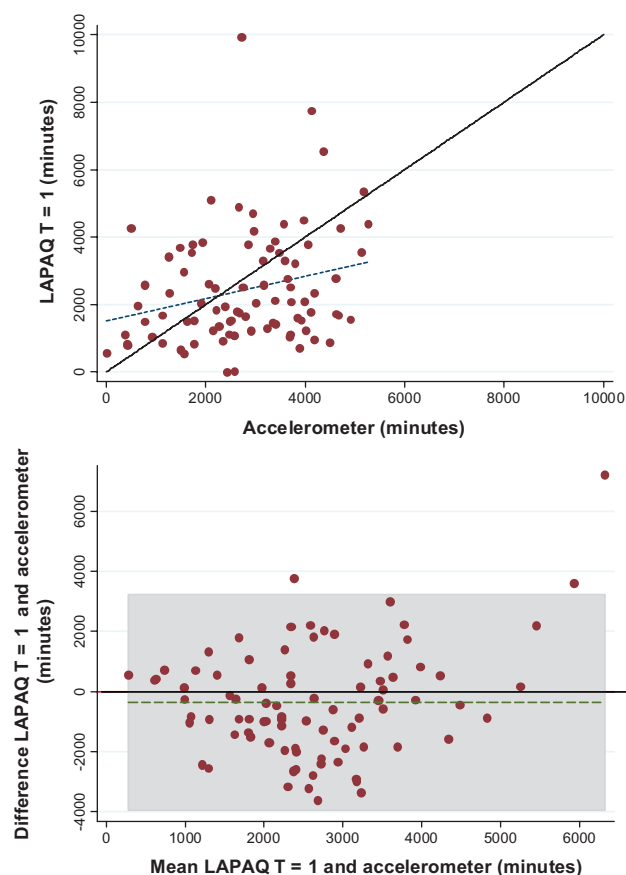


Figure 1 Scatter plot (A) and Bland–Altman plot (B).

Notes: Solid line indicates the reference line (scatter plot: $x = y$, and Bland–Altman plot: $y = 0$). Dotted line indicates the regression line. Marked field signifies limits of agreement (± 2 SD). Bland–Altman plot: mean = (LAPAQ T = 1 plus LAPAQ T = 2)/2, difference = LAPAQ T = 1 minus LAPAQ T = 2.

Abbreviations: LAPAQ, Longitudinal Ageing Study Amsterdam Physical Activity Questionnaire; T = 1, 2 weeks; T = 2, 4 weeks.

the LAPAQ was administered twice, but only within a 2-week interval. This probably avoided memory effects in participants' ability to report their level of PA, while at the same time administration of the questionnaire in this timeframe may have ensured stability of activity levels. A drawback of wearing the ACTR for only the first 2 weeks is that it may have induced a Hawthorne effect, affecting our reproducibility measurements. A third strength of our study is that our population consisted of 86 persons who varied in terms of sex, age, BMI, and comorbidities, which enhances the robustness of our correlation measures.

Neilson et al evaluated several validation studies of PA questionnaires and the largest study they found contained a total of 80 subjects (women only).⁹ They also found that in only 4 out of 36 studies, the PA questionnaires covered exactly the same time period of activity as the validation instrument (which was consistent with the second strength of our study). Finally, a fourth strength of our study is the elaborate statis-

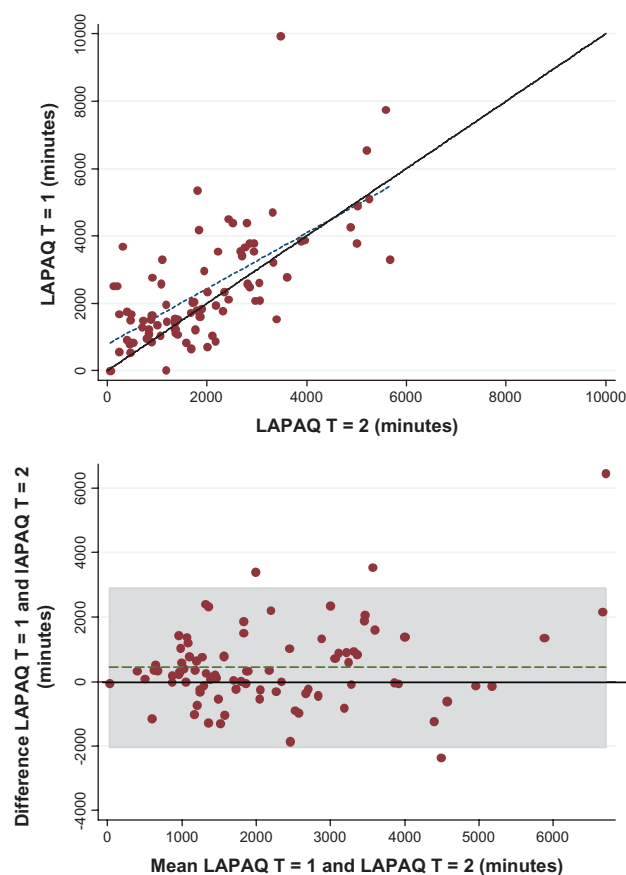


Figure 2 Scatter plot (A) and Bland–Altman plot (B).

Notes: Solid line indicates the reference line (scatter plot: $x = y$, and Bland–Altman plot: $y = 0$). Dotted line indicates the regression line. Marked field signifies limits of agreement (± 2 SD). Bland–Altman plot: mean = (LAPAQ T = 1 plus accelerometer)/2, difference = LAPAQ T = 1 minus accelerometer. Note that, in the scatterplot, subjects above the reference line overestimated their duration of PA on the LAPAQ compared to the accelerometer measurements, while subjects below this line underestimated their duration of PA on the LAPAQ.

Abbreviations: T = 1, 2 weeks; T = 2, 4 weeks; LAPAQ, Longitudinal Ageing Study Amsterdam Physical Activity Questionnaire; PA, physical activity.

tical approach. Ambiguity still exists about the appropriate statistical methods and interpretation of validation studies. In Schmidt and Steindorf's systematic review,²² the most common approach in validation studies is still the presentation of correlation coefficients (41 of 46 articles). However, the limitations of correlation coefficients are well documented in the literature.^{22,23} The appropriate evaluation methods recommended by Bland and Altman were only applied in 10/46 publications. Schmidt and Steindorf showed that serious bias in questionnaires can be revealed by Bland–Altman plots, but they may remain undetected by correlation coefficients. In our study, correlation coefficients were added as a point of comparison for previous studies, but our primary interest was to compare our results by graphically visualizing aspects of agreement in Bland–Altman plots. A limitation of our study (which also served as a strength) is the limited sample size.

Although the present study seems to be one of the largest studies in this field, we still face wide confidence intervals indicating a high degree of variability in our findings.

In conclusion, the LAPAQ seems unsuitable for exact measurement of physical activity levels in older adults. However, the LAPAQ is a fast and practical self-administered questionnaire that can be used in practice and in studies to determine if a person's activity level is above the recommendation level of the ACSM and the AHA.

Acknowledgments

The authors regret to announce the passing of Leo Beem prior to publication of this paper. We would like to thank all participants and GPs from the Klein Gooi-Oord health care center. Also, thanks go to Alice Karsten and Albertien Buijs (Academic Medical Center, Department of General Practice, University of Amsterdam, the Netherlands) for their logistical assistance, and to Vianda Stel (Department of Clinical Informatics, Academic Medical Center, University of Amsterdam, the Netherlands) for advice on this study. This work was supported by the Dutch Asthma Foundation (grant # 3.4.07.045).

Authors' contributions

LS, SW, and GR drafted and revised the manuscript. All authors read and corrected draft versions of the manuscript and approved the final manuscript.

Disclosure

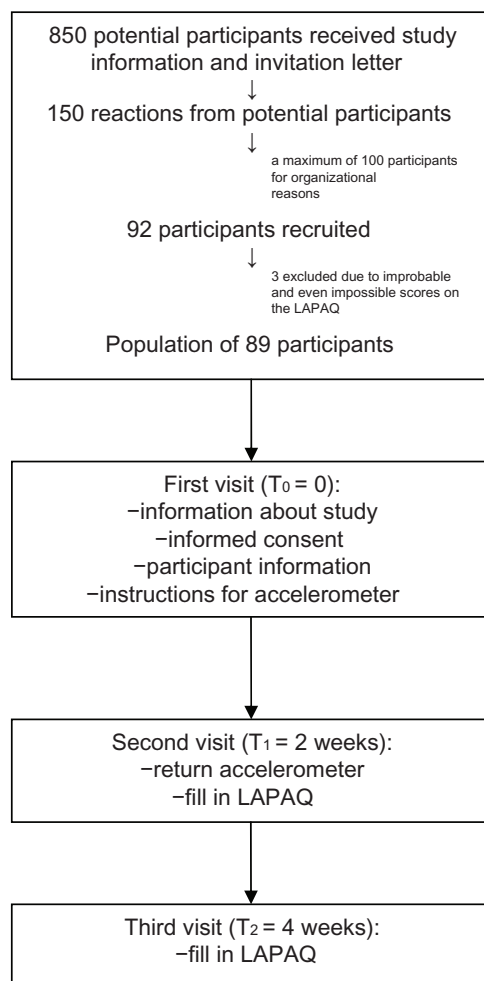
The authors report no conflicts of interest in this work.

References

- Garcia-Aymerich J, Lange P, Benet M, Schnohr P, Anto JM. Regular physical activity reduces hospital admission and mortality in chronic obstructive pulmonary disease: a population based cohort study. *Thorax*. 2006;61(9):772–778.
- Garcia-Aymerich J, Lange P, Benet M, Schnohr P, Anto JM. Regular physical activity modifies smoking-related lung function decline and reduces risk of chronic obstructive pulmonary disease: a population-based cohort study. *Am J Respir Crit Care Med*. 2007;175(5):458–463.
- Garcia-Aymerich J, Serra I, Gomez FP, et al; for Phenotype and Course of COPD Study Group. Physical activity and clinical and functional status in COPD. *Chest*. 2009;136(1):62–70.
- Hillsdon M, Foster C, Thorogood M. Interventions for promoting physical activity. *Cochrane Database Syst Rev*. 2005;1:CD003180.
- Ruigomez A, Alonso J, Anto JM. Relationship of health behaviours to five-year mortality in an elderly cohort. *Age Ageing*. 1995;24(2):113–119.
- Warburton DE, Charlesworth S, Ivey A, Nettlefold L, Bredin SS. A systematic review of the evidence for Canada's Physical Activity Guidelines for Adults. *Int J Behav Nutr Phys Act*. 2010;7:39.
- Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. *CMAJ*. 2006;174(6):801–809.
- ACSM News Releases [homepage on the Internet]. ASCM issues new recommendations on quantity and quality of exercise. Indianapolis: American College of Sports Medicine; updated 2008. Available from: <http://www.acsm.org/about-acsm/media-room/news-releases/2011/08/01/acsm-issues-new-recommendations-on-quantity-and-quality-of-exercise>. Accessed 21 June, 2012.
- Neilson HK, Robson PJ, Friedenreich CM, Csizmadia I. Estimating activity energy expenditure: how valid are physical activity questionnaires? *Am J Clin Nutr*. 2008;87(2):279–291.
- Melanson EL Jr, Freedson PS. Physical activity assessment: a review of methods. *Crit Rev Food Sci Nutr*. 1996;36(5):385–396.
- Lamonte MJ, Ainsworth BE. Quantifying energy expenditure and physical activity in the context of dose response. *Med Sci Sports Exerc*. 2001;33(Suppl 6):S370–S378.
- Harris TJ, Owen CG, Victor CR, Adams R, Ekelund U, Cook DG. A comparison of questionnaire, accelerometer, and pedometer: measures in older people. *Med Sci Sports Exerc*. 2009;41(7):1392–1402.
- Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am J Clin Nutr*. 1982;36(5):936–942.
- Craig CL, Marshall AL, Sjoström M, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*. 2003;35(8):1381–1395.
- Stel VS, Smit JH, Pluijm SM, Visser M, Deeg DJ, Lips P. Comparison of the LASA Physical Activity Questionnaire with a 7-day diary and pedometer. *J Clin Epidemiol*. 2004;57(3):252–258.
- Andre D, Pelletier R, Farringdon J, et al. The development of the SenseWear® armband, a revolutionary energy assessment device to assess physical activity and lifestyle. BodyMedia, Inc; 2006. Available from: http://www.bodybugg.com/pdf/wp_accuracy_ee.pdf. Accessed 21 June, 2012.
- King GA, Torres N, Potter C, Brooks TJ, Coleman KJ. Comparison of activity monitors to estimate energy cost of treadmill exercise. *Med Sci Sports Exerc*. 2004;36(7):1244–1251.
- St-Onge M, Mignault D, Allison DB, Rabasa-Lhoret R. Evaluation of a portable device to measure daily energy expenditure in free-living adults. *Am J Clin Nutr*. 2007;85(3):742–749.
- Wadsworth D, Howard T, Hallam JS, Blunt G. A validation study of a continuous body-monitoring device: assessing energy expenditure at rest and during exercise. *Med Sci Sports Exerc*. 2005;37(5):S24.
- Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of physical activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc*. 1993;25(1):71–80.
- Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc*. 2000;32(Suppl 9):S498–S504.
- Schmidt ME, Steindorf K. Statistical methods for the validation of questionnaires – discrepancy between theory and practice. *Methods Inf Med*. 2006;45(4):409–413.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–310.
- Fletcher R, Fletcher S. Chapter 8 Treatment. In: *Clinical Epidemiology: The Essentials*. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2005:131.

Appendices

Appendix I



Appendix I Flow chart of the study.

Abbreviations: LAPAQ, Longitudinal Ageing Study Amsterdam Physical Activity Questionnaire; T, time.

Appendix 2 LAPAQ

1. Do you walk outside?

Explanation: with walking outside we mean walking to go shopping or doing other daily activities, like visiting someone.

We do not mean: a walking tour.

☐ No (go to question 4)

☐ Yes

2. How many times did you walk during the past two weeks?

|_|_| Times in the past two weeks

3. How long did you usually walk each time?

|_|_| Hours and |_|_| Minutes

4. Do you cycle?

Explanation: with cycling we mean cycling to go shopping or doing other daily activities, like visiting someone. With cycling we do not mean: a cycling tour.

☐ No (go to question 7)

☐ Yes

5. How many times did you cycle during the past two weeks?

|_|_| Times in the past two weeks

6. How long did you usually cycle each time?

|_|_| Hours and |_|_| Minutes

7. Do you have a garden (including allotment)?

☐ No (go to question 10)

☐ Yes

8. How many times did you work in the garden during the past two weeks?

|_|_| Times in the past two weeks

9. How long did you usually work in the garden each time?

|_|_| Hours and |_|_| Minutes

10. Do you do sports?

☐ No (go to question 12)

☐ Yes

11. Which sports did you do during the past two weeks?

Mark the sports that you did do during the past two weeks and fill in how many times and for how long each time you did those sports.

SPORTS	Times in the past two weeks	Duration each time
<input type="checkbox"/> 1. Gymnastics	_ _	_ _ Hours and _ _ Min
<input type="checkbox"/> 2. Cycling on hometrainer	_ _	_ _ Hours and _ _ Min
<input type="checkbox"/> 3. Distance cycling	_ _	_ _ Hours and _ _ Min
<input type="checkbox"/> 4. Distance walking	_ _	_ _ Hours and _ _ Min
<input type="checkbox"/> 5. Swimming	_ _	_ _ Hours and _ _ Min
<input type="checkbox"/> 6. Tennis/Badminton	_ _	_ _ Hours and _ _ Min
<input type="checkbox"/> 7. Winter sports	_ _	_ _ Hours and _ _ Min
<input type="checkbox"/> 8. Cardio-fitness	_ _	_ _ Hours and _ _ Min
<input type="checkbox"/> 9. Other	_ _	_ _ Hours and _ _ Min

12. Do you do light household tasks?

Explanation: with light household tasks we mean washing the dishes, dusting, making the bed, doing the laundry, hanging out the laundry, ironing, tidying up and cooking meals.

☐ No (go to question 15)

☐ Yes

13. How many days did you do light household tasks during the past two weeks?

|_|_| days in the past two weeks

14. How long did you usually do light household tasks each day?

|_|_| Hours and |_|_| Minutes

15. Do you do heavy household tasks?

Explanation: with heavy household tasks we mean window cleaning, changing the bed, beating the mat, vacuuming, washing or scrubbing the floor and chores with sawing, carpeting, repairing or painting.

☐ No (go to question 18)

☐ Yes

16. How many days did you do heavy household tasks during the past two weeks?

|_|_| days in the past two weeks

17. How long did you usually do heavy household tasks each day?

|_|_| Hours and |_|_| Minutes

18. You just answered questions about your usual activities of the past two weeks.

Were the past two weeks normal as compared to the rest of the past year?

☐ Yes (end of questionnaire)

☐ No, I did more because

☐ good weather

☐ bad weather

☐ holiday

☐ other:

☐ No, I did less because

☐ good weather

☐ bad weather

☐ holiday

☐ illness

☐ other:

Clinical Epidemiology

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification, systematic

Submit your manuscript here: <http://www.dovepress.com/clinical-epidemiology-journal>

reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Dovepress